

Result	Score	Query	Match	Length	DB	ID	Description
No.		#					
1	2719	100.0	2719	9	US-09-866-8668-9	Sequence 60, Appl	
2	2700.6	99.3	2718	15	US-10-120-6687-60	Sequence 69, Appl	
3	2662.6	97.9	2883	15	US-10-101-510-639	Sequence 639, Appl	
4	2354.6	86.6	2574	9	US-09-981-353-34	Sequence 34, Appl	
5	2352.8	86.5	2418	10	US-09-961-088-2	Sequence 2, Appl	
6	2221.4	81.7	2247	9	US-09-866-8668-26	Sequence 26, Appl	
7	2090.4	76.9	2788	9	US-09-745-763-196	Sequence 196, Appl	
8	2022.2	73.4	2027	16	US-10-405-806-1	Sequence 1, Appl	
9	2010	73.9	2053	16	US-10-405-806-12	Sequence 13, Appl	
10	1392.2	51.2	2025	9	US-09-866-8668-13	Sequence 13, Appl	
11	746.4	27.5	795	10	US-09-961-086-7	Sequence 7, Appl	
12	445	16.4	445	13	US-10-052-288-201	Sequence 201, Appl	
13	395.8	14.6	427	9	US-09-860-352-11649	Sequence 11649, Appl	
14	364	13.4	427	9	US-09-860-352-12839	Sequence 12839, A	

15	355.4	13.1	467	US-09-960-355-9050	Sequence 9050, Ap
16	351.6	12.9	414	US-09-960-355-5514	Sequence 5514, Ap
17	351.6	12.9	465	US-09-983-965-1505	Sequence 1505, Ap
18	347.4	12.8	434	US-10-313-669-98	Sequence 86, App
19	347.4	12.8	456	US-09-983-965-1396	Sequence 1396, Ap
20	333.2	12.2	502	US-09-866-866A-11	Sequence 11,3, Ap1
21	332.3	12.2	410	US-09-983-965-1133	Sequence 11,33, Ap
22	331.2	12.2	408	US-09-983-965-1322	Sequence 1232, Ap
23	330.4	12.2	377	US-09-960-355-4326	Sequence 4326, Ap
24	329	12.1	421	US-09-983-965-1800	Sequence 1800, Ap
25	327.2	12.1	410	US-09-983-965-1441	Sequence 1441, Ap
26	321.6	11.8	410	US-09-983-965-1511	Sequence 1511, Ap
27	320.8	11.8	387	US-09-960-352-6470	Sequence 6470, Ap
28	313.8	11.5	419	US-09-983-965-1347	Sequence 1247, Ap
29	313.2	11.5	422	US-09-960-355-3985	Sequence 3985, Ap
30	297.2	10.9	442	US-09-983-965-1765	Sequence 1765, Ap
31	297	10.9	441	US-09-983-965-1495	Sequence 1495, Ap
32	293.4	10.8	367	US-09-983-965-1118	Sequence 1218, Ap
33	289.2	10.6	366	US-09-960-352-4458	Sequence 4458, Ap
34	287.8	10.6	406	US-09-983-965-1308	Sequence 1308, Ap
35	286.2	10.5	419	US-09-983-965-1431	Sequence 1431, Ap
36	285	10.5	390	US-09-983-965-1351	Sequence 1351, Ap
37	285.8	10.5	418	US-09-983-965-1196	Sequence 1196, Ap
38	282.8	10.4	415	US-09-983-965-1435	Sequence 1435, Ap
39	279	10.3	451	US-09-983-965-1658	Sequence 1358, Ap
40	278	10.2	414	US-09-983-965-1311	Sequence 1311, Ap
41	271.4	10.0	384	US-09-983-965-1188	Sequence 1188, Ap
42	269.4	9.9	399	US-09-983-965-1476	Sequence 1476, Ap
43	268.4	9.9	398	US-09-983-965-1310	Sequence 1310, Ap
44	267.6	9.8	499	US-09-833-381-965	Sequence 965, App
45	261.4	9.6	391	US-09-983-965-1108	Sequence 1108, Ap

## ALIGNMENTS

```

: RESULT 1
: US-09-866-866A-9
: Sequence 9, Application US/09866866A
: Patent No. US20020102244A1
: GENERAL INFORMATION:
: APPLICANT: Sorrentino, Brian
: APPLICANT: Schuetz, John
: TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
: FILE REFERENCE: 1340-1-021CIP2
: CURRENT APPLICATION NUMBER: US/09/866, 866A
: CURRENT FILING DATE: 2001-08-30
: PRIOR APPLICATION NUMBER: 09/584, 586
: PRIOR FILING DATE: 2000-05-31
: PRIOR APPLICATION NUMBER: PCT/US99/11825
: PRIOR FILING DATE: 1999-05-27
: PRIOR FILING DATE: 1998-05-28
: PRIOR FILING DATE: 1998-05-28
: NUMBER OF SEQ ID NOS: 27
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 9
: LENGTH: 2719
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-866-866A-9

```

Query Match	100.0%;	Score 2719;	DB 9;	Length 2719;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2719; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

Qy 1 TTATGAAAGGCAACCGGCAATGCTGGAGGCTTTGTAATGGAATAATGCTGCTTAGA 60  
Db 1 TTATGAAAGGCAACCGGCAATGCTGGAGGCTTTGTAATGGAATAATGCTGCTTAGA 60  
Qy 61 GTTTGTTTGAAGAGTCCGGGTGACTGATCCCAACATTACATCCTTAATTTGTTAAAGGC 120  
Db 61 GTTTGTTTGAAGAGTCCGGGTGACTGATCCCAACATTACATCCTTAATTTGTTAAAGGC 120

QY	12	GGCGTCGAGGCGACGCACTCGTGAGATCGTAGGCTTGTTAAGCCGACCTCATATTA	180
Db	121	TGCCTCGAGGCGACGCACTCGTGAGATCGTAGGCTTGTTAAGCCGACCTCATATTA	180
QY	181	GCTGAAAAGATPAAAACCTTCGAGATGCTTCCAGTAATGTGCAAGTTTATCCCAAGT	240
Db	181	GCTGAAAAGATPAAAACCTTCGAGATGCTTCCAGTAATGTGCAAGTTTATCCCAAGT	240
QY	241	TCACAGAAGAAACACCAATGGCTGCCCGGCGACAGTTTCATGACCTGAAAGCATTTACT	300
Db	241	TCACAGAAGAAACACCAATGGCTGCCCGGCGACAGTTTCATGACCTGAAAGCATTTACT	300
QY	301	GAAAGAGCTGTGTAGTTTCAATPACATCTGCTACATGAGTAAACCTGAAAGTGGCTTT	360
Db	301	GAAAGAGCTGTGTAGTTTCAATPACATCTGCTACATGAGTAAACCTGAAAGTGGCTTT	360
QY	361	CTACCTTGCGAAAAACGAGTTGAGAAAAATATATGATGATPACATGAGGATCATGAA	420
Db	361	CTACCTTGCGAAAAACGAGTTGAGAAAAATATATGATGATPACATGAGGATCATGAA	420
QY	421	CTGTGTCGAAAGCCATCTCTGGACCCACAGGTGAGGCAAACTTCGTTATTTGATGTC	480
Db	421	CTGTGTCGAAAGCCATCTCTGGACCCACAGGTGAGGCAAACTTCGTTATTTGATGTC	480
QY	481	TTAGCGTCAGAAAAGATCCAAAGTGATTAATCTGGAGATGTTCTGATTAATGAGACCG	540
Db	481	TTAGCGTCAGAAAAGATCCAAAGTGATTAATCTGGAGATGTTCTGATTAATGAGACCG	540
QY	541	CGACCTGCGAAATTTCAAAATGTAAATCAAGTTACGTGTACAAAGATGATTTGTGATGGC	600
Db	541	CGACCTGCGAAATTTCAAAATGTAAATCAAGTTACGTGTACAAAGATGATTTGTGATGGC	600
QY	601	ACTCGACGAGAGAGAAACCTTACAGTCTCAGACAGCTCTCGGCTGCAACAATATG	660
Db	601	ACTCGACGAGAGAGAAACCTTACAGTCTCAGACAGCTCTCGGCTGCAACAATATG	660
QY	661	ACGAATCATGAAAAAACGAAACGATTAACAGGGCTCATTGAAGATTAGTCTGATPAA	720
Db	661	ACGAATCATGAAAAAACGAAACGATTAACAGGGCTCATTGAAGATTAGTCTGATPAA	720
QY	721	GTGCGAGACTCCAAAGTTGAACTCAGTTTATCCGTGTGTGCTCGAGAGAAAGAAA	780
Db	721	GTGCGAGACTCCAAAGTTGAACTCAGTTTATCCGTGTGTGCTCGAGAGAAAGAAA	780
QY	781	AGGACTAGTANAGAAATGAGCTTATCACTGATCCTTCATCTGTGCTTGGATGAGCT	840
Db	781	AGGACTAGTANAGAAATGAGCTTATCACTGATCCTTCATCTGTGCTTGGATGAGCT	840
QY	841	ACAACTGCTTAGACTCAAGCAGACCAATGTGCTCTTTGCTCCTGAAAAAGATGTCT	900
Db	841	ACAACTGCTTAGACTCAAGCAGACCAATGTGCTCTTTGCTCCTGAAAAAGATGTCT	900
QY	901	AAGCAGGACGCAACATCATCTTCTCCATTCATCAGCCTCGATATTCATCTTCAAGTG	960
Db	901	AAGCAGGACGCAACATCATCTTCTCCATTCATCAGCCTCGATATTCATCTTCAAGTG	960
QY	961	TTGTAGACCTTACCTTTATGGCTCCAGGAAAGCTTATGTTCCAGGGCCTGCTCAGAG	1020
Db	961	TTGTAGACCTTACCTTTATGGCTCCAGGAAAGCTTATGTTCCAGGGCCTGCTCAGAG	1020
QY	1021	GCTTGGGATPCTTTGAAATCAGCTGTTATCACTGAGGCTTAATPAAACCTGCAAG	1080
Db	1021	GCTTGGGATPCTTTGAAATCAGCTGTTATCACTGAGGCTTAATPAAACCTGCAAG	1080
QY	1081	TTCTTCTTGGACATCATTAATGAGATTTCACTGCTGTGGCATTTAAACAGAGAAAGAC	1140
Db	1081	TTCTTCTTGGACATCATTAATGAGATTTCACTGCTGTGGCATTTAAACAGAGAAAGAC	1140
QY	1141	TTTAAAGCACAGAGATCATPAGGCTTCCAAAGCAAGATPAAAGCCATCATGAAATTA	1200
Db	1141	TTTAAAGCACAGAGATCATPAGGCTTCCAAAGCAAGATPAAAGCCATCATGAAATTA	1200
QY	1201	GGGAGATTATGTCACTCTCTCTTCTAACAAAGACAAAAGCTGAATTACATCACTT	1266

Db	1201	GC	GGAGATTTATGTCAACTCCTCCTTCTACAAAGAGC	1260
Qy	1261	TC	CGGGGTGAGAAAGAAAGATACAGTCTTCAAGAGATCAGCTACACCACTCC	1320
Db	1261	TC	CGGGGGGTGAGAAAGAAAGATACAGTCTTCAAGAGATCAGCTACACCACTCC	1320
Qy	1321	TT	CGTATCAACATCAGATGGGTTC	1380
Db	1321	TT	CGTATCAACATCAGATGGGTTC	1380
Qy	1381	CAG	CCCTCTATAGTCAATCTTGTACAGTGTATCGGAGCTGGTATAGGGCCATT	1440
Db	1381	CAG	CCCTCTATAGTCAATCTTGTACAGTGTATCGGAGCTGGTATAGGGCCATT	1440
Qy	1441	TAC	TTTGGGCTAA	1500
Db	1441	TAC	TTTGGGCTAA	1500
Qy	1501	CT	GCACCAACCAAGTCTTTCAGAGTGTTCAGCCGTGGA	1560
Db	1501	CT	GCACCAACCAAGTCTTTCAGAGTGTTCAGCCGTGGA	1560
Qy	1561	AAG	CTCTTCATACATGATACATCAGCGGATACACAGAGTGCATCTTATTCCTTGA	1620
Db	1561	AAG	CTCTCTTCATACATGATACATCAGCGGATACACAGAGTGCATCTTATTCCTTGA	1620
Qy	1621	AA	CTGTATCTATCTATTTATTTATCCATAGAGATGTTCACCAAGTATATTTATTCCTGTATA	1680
Db	1621	AA	CTGTATCTATCTATTTATTTATCCATAGAGATGTTCACCAAGTATATTTATTCCTGTATA	1680
Qy	1681	GT	GTACTTCAATGTAGGATGGAAGC	1740
Db	1681	GT	GTACTTCAATGTAGGATGGAAGC	1740
Qy	1741	CT	TATGATGTGGCTTATTCAGCCAGTTCATGGCACTTGCCATAGAGACAGACGTCAAGT	1800
Db	1741	CT	TATGATGTGGCTTATTCAGCCAGTTCATGGCACTTGCCATAGAGACAGACGTCAAGT	1800
Qy	1801	GT	GTGTCTCTGAGCAACACTTCTCATGACCACTGTGTTTGTTTATGATGATTTTTC	1860
Db	1801	GT	GTGTCTCTGAGCAACACTTCTCATGACCACTGTGTTTGTTTATGATGATTTTTC	1860
Qy	1861	GG	CTGTGTGGCATCTCCACAACATGGATCTTGGGTCAGGCTCAGTACCTCAGC	1920
Db	1861	GG	CTGTGTGTGGCATCTCCACAACATGGATCTTGGGTCAGGCTCAGTACCTCAGC	1920
Qy	1921	AT	TCACAGATATGAGATTTACGGCTTGCAGCATATGAAATTTTGGGACAAACCTTCTGC	1980
Db	1921	AT	TCACACAGATATGAGATTTACGGCTTGCAGCATATGAAATTTTGGGACAAACCTTCTGC	1980
Qy	1981	CA	GAGATCATGGAACAGAAACAACTCTGTATCATGGAATGATGACATGGGGAAGAA	2040
Db	1981	CA	GAGATCATGGAACAGAAACAACTCTGTATCATGGAATGATGACATGGGGAAGAA	2040
Qy	2041	TAT	TGTGTAAGCAGGGATGATCTTCACCTGGGGCTTGTGAAGATCACGTGGCC	2100
Db	2041	TAT	TGTGTGTAAGCAGGGATGATCTTCACCTGGGGCTTGTGAAGATCACGTGGCC	2100
Qy	2101	TT	GGCTGTATGATGATTTATTCCTCACAATTCCTACCTGGAATTTTATTTCTTTAA	2160
Db	2101	TT	GGCTGTATGATGATTTATTTCTCACAATTCCTACCTGGAATTTTATTTCTTTAA	2160
Qy	2161	AA	ATATTTCTTAAATTTCCCTTAAATTCAGTATGATTTATCTCTCAATPAAAAAGAGAC	2220
Db	2161	AA	ATATATTTCTTAAATTTCCCTTAAATTCAGTATGATTTATCTCTCAATPAAAAAGAGAC	2220
Qy	2221	TT	GTATGATGATTCATCAAGTTTTTTTGTATTTCTGTCCCTGTGCATCACACTG	2280
Db	2221	TT	GTATGATGATTCATCAAGTTTTTTTGTATTTCTGTCCCTGTGCATCACACTG	2280
Qy	2281	TT	GCACAGACCAATGTTTTTAAAGATACATTTTGAATATCAACAACACTGATTA	2340

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 18, 2004, 11:11:46 ; Search time 721 Seconds

(without alignments)  
16020.597 Million cell updates/sec

Title: US-09-856-927-1

Perfect score: 2719

Sequence: 1 tttagaacgcacgcgtgcac.....catgaagtaataagaactt 2719

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues  
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
1: geneseq1808:\*  
2: geneseq1808:\*  
3: geneseq20006:\*  
4: geneseq2001as:\*  
5: geneseq2001bs:\*  
6: geneseq2002as:\*  
7: geneseq2003as:\*  
8: geneseq2003bs:\*  
9: geneseq2003cs:\*  
10: geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2719	100.0	2719	3	AA294760 Human ATP
2	2719	100.0	2719	3	AA27938 ATP-Bind
3	2719	100.0	2719	6	ABA4365 Human BCR
4	2700.6	99.3	2718	7	ACC80605 Human ABC
5	2662.6	97.9	2883	6	AB255528 Human gen
6	2354.6	86.6	2574	4	AA27724 Human tra
7	2354.6	86.6	2574	8	ADAI0916 Human CDN
8	2352.8	86.5	2418	2	AA206360 Breast Ca
9	2321.4	81.7	2247	6	ABA4383 Human BCR
10	2090.4	76.9	2788	2	AAV55745 Human sec
11	2090.4	76.9	2788	6	ABQ2071 Human pol
12	2044.2	75.2	2077	4	AAH15008 Human CDN
13	2022.2	74.4	2053	6	ABK49901 CDNA enco
14	2010	73.9	2053	6	ABK49911 CDNA enco
15	1963.2	72.2	1968	9	AD54181 Human bre
16	1963.2	72.2	1968	9	AD54181 Human bre
17	1960	72.1	1998	6	AA142414 Human BCR
18	1955.2	71.9	1998	6	AA142413 Human BCR
19	1932.2	51.2	2025	6	ABA4371 Murine BC
20	746.4	27.5	795	2	AA206365 BCRP frag
21	679.8	25.0	727	4	AAH07859 Human CDN
22	501.4	18.4	594	4	AAH12829 Human CDN
23	461.4	17.0	565	7	AB282925 Toxicolog

24	450.8	16.6	707	3	AA01965 Human col
25	445	16.4	445	5	AA293380 Lung carc
26	395.8	14.6	447	7	ABX46484 Bovine ES
27	364	13.4	427	7	ABX47674 Bovine ES
28	355.4	13.1	467	7	ABX43885 Bovine ES
29	351.6	12.9	414	7	ABX40349 Bovine ES
30	351	12.9	465	7	ABX51576 Bovine ES
31	347	12.8	456	7	ABX51467 Bovine ES
32	333	12.2	502	3	AA27939 ATP-Bind
33	333	12.2	502	6	ABA4370 Mouse BCR
34	332.2	12.2	410	7	ABX51204 Bovine ES
35	331.2	12.2	408	7	ABX51293 Bovine ES
36	330.4	12.2	378	7	ABX39161 Bovine ES
37	329	12.1	421	7	ABX51871 Bovine ES
38	327.2	12.0	404	7	ABX51512 Bovine ES
39	321.6	11.8	410	7	ABX51582 Bovine ES
40	320.8	11.8	387	7	ABX41305 Bovine ES
41	319.8	11.5	419	7	ABX51318 Bovine ES
42	313.2	11.5	422	7	ABX38820 Bovine ES
43	297.2	10.9	442	7	ABX51836 Bovine ES
44	297	10.9	441	7	ABX51566 Bovine ES
45	293.4	10.8	367	7	ABX51289 Bovine ES

## ALIGNMENTS

RESULT 1	AA294760	standard; CDNA; 2719 BP.
XX	AA294760	
AC	AA294760	
DT	01-AUG-2000	(first entry)
XX		
DE	Human ATP binding cassette ABCG2 CDNA.	
XX		
KM	ABCG2; ATP binding cassette; human; cholesterol; lipid disorder; atherosclerosis; lipid disorder; dyslipidemia; peristalsis;	
KM	lupus erythematosus; diagnosis; gene therapy; WRP4;	
KM	multidrug resistance associated protein; chromosome 4q22-23; ss.	
OS	Homo sapiens.	
PN	W0200018912-R2.	
XX		
PD	06-APR-2000.	
XX		
PF	21-SEP-1999; 99WO-EP006991.	
XX		
PR	25-SEP-1998; 98US-0101706P.	
XX		
PA	(PARB) BAYER AG.	
PI	Schmitz G, Klucken J;	
DR	WPI; 2000-293151/25.	
XX		
PT	Adenosine triphosphate binding proteins useful for identifying agents for treating atherosclerosis and other inflammatory disorders.	
XX		
PS	Claim 9; Page 138-139; 154pp; English.	

The present sequence is that of human ATP binding cassette subfamily C protein ABCG2 CDNA. The CDNA was identified using a differential display method in which monocytes from peripheral blood were subjected to macrophage differentiation and cholesterol loading with acetylated low density lipoproteins and subsequent deacetylating with high density lipoprotein (HDL) to identify cholesterol sensitive genes. The gene maps to chromosome 4q22-23. The invention provides cholesterol-sensitive ABC genes (see AA294734-63). These genes, and polypeptides encoded by them, can be used for diagnostic and therapeutic applications, and for biochemical or cell-based assays to screen for pharmacologically active

modulator compounds useful for the treatment of lipid disorders,  
atherosclerosis or other inflammatory diseases such as psoriasis and  
lupus erythematosus

Sequence 2719 BP: 799 A; 545 C; 564 G; 811 T; 0 U; 0 Other;  
Query Match 100.0%; Score 2719; DB 3; Length 2719;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2719; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TTTGGAACGACCGTGACATGCTGGTGGTCTTGTAGTGAAGAACTGCTGTAGA 60  
1 TTTGGAACGACCGTGACATGCTGGTGGTCTTGTAGTGAAGAACTGCTGTAGA 60  
61 GTTGTGGAAGGTCGGGTGACTCATCCCAATTTTAACTCTTAAATGTTAAAGCG 120  
61 GTTGTGGAAGGTCGGGTGACTCATCCCAATTTTAACTCTTAAATGTTAAAGCG 120  
121 TGCCCTCCGAGCGACGACATCTCTGAGATCTCTGAGCCCTTGTAGACCGACTATTA 180  
121 TGCCCTCCGAGCGACGACATCTCTGAGATCTCTGAGCCCTTGTAGACCGACTATTA 180  
121 TGCCCTCCGAGCGACGACATCTCTGAGATCTCTGAGCCCTTGTAGACCGACTATTA 180  
181 GCTGAAAGATTAATAACTCTCCAGATGCTTTCAGTAATGTCGAAGTTTATCCCAAG 240  
181 GCTGAAAGATTAATAACTCTCCAGATGCTTTCAGTAATGTCGAAGTTTATCCCAAG 240  
181 GCTGAAAGATTAATAACTCTCCAGATGCTTTCAGTAATGTCGAAGTTTATCCCAAG 240  
241 TCACAGGAAACACCAATGCTTCCCGCGACAGTTTCCATGATCTGAAAGCCTTACT 300  
241 TCACAGGAAACACCAATGCTTCCCGCGACAGTTTCCATGATCTGAAAGCCTTACT 300  
241 TCACAGGAAACACCAATGCTTCCCGCGACAGTTTCCATGATCTGAAAGCCTTACT 300  
301 GAAAGAGCTGTGTTAAGTTTCAATCATCTGCTATCAGTAATACTGAAAGAGGCTTT 360  
301 GAAAGAGCTGTGTTAAGTTTCAATCATCTGCTATCAGTAATACTGAAAGAGGCTTT 360  
301 GAAAGAGCTGTGTTAAGTTTCAATCATCTGCTATCAGTAATACTGAAAGAGGCTTT 360  
361 CTACCTTGTGAAAACCAAGTTGAGAAAGAAATTTATGAAATATCAATGGATCATGAA 420  
361 CTACCTTGTGAAAACCAAGTTGAGAAAGAAATTTATGAAATATCAATGGATCATGAA 420  
421 CCGTGTTCACAGCCATCTGAGGACCCACAGGTGAGGCAATCTGCTATTTAGATGTC 480  
421 CCGTGTTCACAGCCATCTGAGGACCCACAGGTGAGGCAATCTGCTATTTAGATGTC 480  
421 CCGTGTTCACAGCCATCTGAGGACCCACAGGTGAGGCAATCTGCTATTTAGATGTC 480  
481 TTAGTGCAGAGAAAGATCCAGATGATTTATCTGAGATGTTCTGATTAATGAGCACCG 540  
481 TTAGTGCAGAGAAAGATCCAGATGATTTATCTGAGATGTTCTGATTAATGAGCACCG 540  
481 TTAGTGCAGAGAAAGATCCAGATGATTTATCTGAGATGTTCTGATTAATGAGCACCG 540  
541 CGACCTGCGCAATTTCAAAATGTAATCAGGTTACGCTGACAAAGATGTTGATGGGC 600  
541 CGACCTGCGCAATTTCAAAATGTAATCAGGTTACGCTGACAAAGATGTTGATGGGC 600  
541 CGACCTGCGCAATTTCAAAATGTAATCAGGTTACGCTGACAAAGATGTTGATGGGC 600  
601 ACTGTCAGGTGAGAGAAACCTTACAGTTCTCAGACAGCTTCCGCTTGCAACAATATG 660  
601 ACTGTCAGGTGAGAGAAACCTTACAGTTCTCAGACAGCTTCCGCTTGCAACAATATG 660  
601 ACTGTCAGGTGAGAGAAACCTTACAGTTCTCAGACAGCTTCCGCTTGCAACAATATG 660  
661 ACGATTCATGAAAAAAGAAAGAGATTAACAGGTCATGTAAGAGTAAAGTCTGATTA 720  
661 ACGATTCATGAAAAAAGAAAGAGATTAACAGGTCATGTAAGAGTAAAGTCTGATTA 720  
661 ACGATTCATGAAAAAAGAAAGAGATTAACAGGTCATGTAAGAGTAAAGTCTGATTA 720  
721 GTGAGAGATCCAGAGTTGAACTCAGTTTATCCGCTGCTGCTGAGAGAGAGAGAAA 780  
721 GTGAGAGATCCAGAGTTGAACTCAGTTTATCCGCTGCTGCTGAGAGAGAGAGAAA 780  
721 GTGAGAGATCCAGAGTTGAACTCAGTTTATCCGCTGCTGCTGAGAGAGAGAGAAA 780  
781 AGGATTCATGAG 840  
781 AGGATTCATGAG 840  
781 AGGATTCATGAG 840  
841 ACAACTGCTTGAAGTCAAGACAGCAATGCTGCTTGTCTGCTGCTGCTGCTGCTGCT 900  
841 ACAACTGCTTGAAGTCAAGACAGCAATGCTGCTTGTCTGCTGCTGCTGCTGCTGCT 900  
841 ACAACTGCTTGAAGTCAAGACAGCAATGCTGCTTGTCTGCTGCTGCTGCTGCTGCT 900  
901 AAG 960  
901 AAG 960

961 TTTGATAGCTCACCTTATTTGGCTCAGAGAGACTTATTTCCAGGCGCTGCTCAGAG 1020  
961 TTTGATAGCTCACCTTATTTGGCTCAGAGAGACTTATTTCCAGGCGCTGCTCAGAG 1020  
1021 GCTTGGGATTTTGGATTCAGCTGCTTATCAGCTGAGGCTTATTAATCCCTGCAAG 1080  
1021 GCTTGGGATTTTGGATTCAGCTGCTTATCAGCTGAGGCTTATTAATCCCTGCAAG 1080  
1021 GCTTGGGATTTTGGATTCAGCTGCTTATCAGCTGAGGCTTATTAATCCCTGCAAG 1080  
1081 TTTCTTCTGACATCATTAATGAGATTTCACTGCTGCTGCTTAAACAGAGAGAGAG 1140  
1081 TTTCTTCTGACATCATTAATGAGATTTCACTGCTGCTGCTTAAACAGAGAGAGAG 1140  
1141 TTTAAAGCCACAGATCATAGAGCTTCCAGCAGATTAAGCCATCATAGAAAATTA 1200  
1141 TTTAAAGCCACAGATCATAGAGCTTCCAGCAGATTAAGCCATCATAGAAAATTA 1200  
1141 TTTAAAGCCACAGATCATAGAGCTTCCAGCAGATTAAGCCATCATAGAAAATTA 1200  
1201 GCGAGATTTATGATCACTCTCTTACAAAAGAGACAAAAGCTGAATTCATCACT 1260  
1201 GCGAGATTTATGATCACTCTCTTACAAAAGAGACAAAAGCTGAATTCATCACT 1260  
1201 GCGAGATTTATGATCACTCTCTTACAAAAGAGACAAAAGCTGAATTCATCACT 1260  
1261 TCCGAGGAGTGAAGAGAGAGAGATCAAGTCTTCAAGAGATCACTCAACACTCC 1320  
1261 TCCGAGGAGTGAAGAGAGAGAGATCAAGTCTTCAAGAGATCACTCAACACTCC 1320  
1261 TCCGAGGAGTGAAGAGAGAGAGATCAAGTCTTCAAGAGATCACTCAACACTCC 1320  
1321 TTTCTGATCACTGATGAGTGGTTCAGAGCTTCAAGCTTCAAAAATCTGCTGGTATCC 1380  
1321 TTTCTGATCACTGATGAGTGGTTCAGAGCTTCAAGCTTCAAAAATCTGCTGGTATCC 1380  
1321 TTTCTGATCACTGATGAGTGGTTCAGAGCTTCAAGCTTCAAAAATCTGCTGGTATCC 1380  
1381 CAGGCTCTATAGCTCAGATCATGTCACAGTGTACTGAGACTGCTGTTATAGTGCAT 1440  
1381 CAGGCTCTATAGCTCAGATCATGTCACAGTGTACTGAGACTGCTGTTATAGTGCAT 1440  
1381 CAGGCTCTATAGCTCAGATCATGTCACAGTGTACTGAGACTGCTGTTATAGTGCAT 1440  
1441 TACTTGGGCTTAAAAATGATTTCTACTGAAATCCAGAAACAGAGCTGGGCTTCTCTTC 1500  
1441 TACTTGGGCTTAAAAATGATTTCTACTGAAATCCAGAAACAGAGCTGGGCTTCTCTTC 1500  
1441 TACTTGGGCTTAAAAATGATTTCTACTGAAATCCAGAAACAGAGCTGGGCTTCTCTTC 1500  
1501 CTGACAGCAACCAAGTGTTCAGAGCTGTTTCAAGCTGTTTCAAGCTGTTTCAAGCT 1560  
1501 CTGACAGCAACCAAGTGTTCAGAGCTGTTTCAAGCTGTTTCAAGCTGTTTCAAGCT 1560  
1501 CTGACAGCAACCAAGTGTTCAGAGCTGTTTCAAGCTGTTTCAAGCTGTTTCAAGCT 1560  
1561 AAGCTTCTATCATGATTAATCAACAGGATCTACAGAGTCAATCTTATCTCTGGA 1620  
1561 AAGCTTCTATCATGATTAATCAACAGGATCTACAGAGTCAATCTTATCTCTGGA 1620  
1561 AAGCTTCTATCATGATTAATCAACAGGATCTACAGAGTCAATCTTATCTCTGGA 1620  
1621 AAACGTTATCTGATTTATTTACCATAGAGATTTACCAAGATTAATTTACTGATA 1680  
1621 AAACGTTATCTGATTTATTTACCATAGAGATTTACCAAGATTAATTTACTGATA 1680  
1621 AAACGTTATCTGATTTATTTACCATAGAGATTTACCAAGATTAATTTACTGATA 1680  
1681 GTGTACTTCACTGATGAGATTTGAAGCCAAAGGAGATGCTTCTGCTTATGATTTACC 1740  
1681 GTGTACTTCACTGATGAGATTTGAAGCCAAAGGAGATGCTTCTGCTTATGATTTACC 1740  
1681 GTGTACTTCACTGATGAGATTTGAAGCCAAAGGAGATGCTTCTGCTTATGATTTACC 1740  
1741 CTTATGATGCTGCTTATCAAGCAGTTCATGACCTGCTGCTGCTGCTGCTGCTGCT 1800  
1741 CTTATGATGCTGCTTATCAAGCAGTTCATGACCTGCTGCTGCTGCTGCTGCTGCT 1800  
1741 CTTATGATGCTGCTTATCAAGCAGTTCATGACCTGCTGCTGCTGCTGCTGCTGCT 1800  
1801 GTGTGTTCTGAGCAACCTTCTCATGACATCTGTTTGTGTTATGATGATTTTGA 1860  
1801 GTGTGTTCTGAGCAACCTTCTCATGACATCTGTTTGTGTTATGATGATTTTGA 1860  
1801 GTGTGTTCTGAGCAACCTTCTCATGACATCTGTTTGTGTTATGATGATTTTGA 1860  
1861 GGTCTGCTGCTGCTTATCAAGCAGTTCATGACCTGCTGCTGCTGCTGCTGCTGCT 1920  
1861 GGTCTGCTGCTGCTTATCAAGCAGTTCATGACCTGCTGCTGCTGCTGCTGCTGCT 1920  
1861 GGTCTGCTGCTGCTTATCAAGCAGTTCATGACCTGCTGCTGCTGCTGCTGCTGCT 1920  
1861 GGTCTGCTGCTGCTTATCAAGCAGTTCATGACCTGCTGCTGCTGCTGCTGCTGCT 1920  
1921 ATTTCAGATGATGATTTACGCTTTCAGCATATGAAATTTTGGGCAAAAATTTCTGC 1980  
1921 ATTTCAGATGATGATTTACGCTTTCAGCATATGAAATTTTGGGCAAAAATTTCTGC 1980  
1921 ATTTCAGATGATGATTTACGCTTTCAGCATATGAAATTTTGGGCAAAAATTTCTGC 1980  
1981 CAGAGCTCAATGCAACAGAGAAACATCTTGTGATTAATGCAACATGCTGCGGAGAA 2040  
1981 CAGAGCTCAATGCAACAGAGAAACATCTTGTGATTAATGCAACATGCTGCGGAGAA 2040  
1981 CAGAGCTCAATGCAACAGAGAAACATCTTGTGATTAATGCAACATGCTGCGGAGAA 2040

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: July 18, 2004, 13:07:51 ; Search time 144 Seconds

(without alignments)  
10478.552 Million cell updates/sec

Title: US-09-856-927-1

Perfect score: 2719

Sequence: 1 tttaaggaacgcacgtgcac.....catlaagtaataagacct 2719

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*  
1: /cgn2\_6/prodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata/2/ina/5A\_COMB.seq:\*  
4: /cgn2\_6/prodata/2/ina/5B\_COMB.seq:\*  
5: /cgn2\_6/prodata/2/ina/5A\_COMB.seq:\*  
6: /cgn2\_6/prodata/2/ina/5B\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2352.8	86.5	2418	4	US-09-245-808-2
2	746.4	27.5	795	4	US-09-245-808-7
3	267.6	9.8	499	4	US-09-833-381-965
4	89.8	3.3	2031	4	US-09-614-912-137
5	86.8	3.2	4159	4	US-09-614-912-139
6	74.6	2.7	1977	4	US-09-614-912-143
7	61.8	2.3	3376	4	US-09-620-312D-918
8	48.6	1.8	1664976	4	US-08-916-421B-1
9	48	1.8	813	4	US-09-134-001C-2724
10	46.2	1.7	1664376	4	US-08-916-421B-1
11	45.4	1.7	7218	1	US-08-232-463-14
12	44.4	1.6	2193	2	US-08-934-481-1
13	44.4	1.6	2193	4	US-09-290-602-1
14	44.4	1.6	6693	4	US-08-961-582-195
15	44.2	1.6	640681	4	US-09-790-988-1
16	43.8	1.6	5852	1	US-07-867-106-2
17	43.4	1.6	747	4	US-09-134-001C-894
18	42.8	1.6	832	4	US-09-621-976-2813
19	42.6	1.6	11050	4	US-08-956-171E-96
20	42.6	1.6	4403765	3	US-09-103-840A-2
21	42.6	1.6	4411529	3	US-09-103-840A-1
22	42.2	1.6	924	4	US-09-543-681A-3732
23	41.8	1.5	1056	4	US-09-543-681A-4166
24	41.8	1.5	2634	4	US-09-534-238-408
25	41.2	1.5	744	4	US-09-134-000C-895
26	41.2	1.5	13188	4	US-08-961-527-70
27	41	1.5	731	1	US-08-451-405A-2

28	40.6	1.5	10182	4	US-08-956-171E-116	Sequence 116, App
29	40.4	1.5	2317	3	US-08-749-522-5	Sequence 5, Appl
30	40.4	1.5	3974	3	US-08-467-504-3	Sequence 3, Appl
31	40.2	1.5	1185	4	US-09-614-912-161	Sequence 161, App
32	40.2	1.5	6669	4	US-10-204-708-6	Sequence 6, Appl
33	39.8	1.5	666	4	US-09-107-532A-576	Sequence 576, App
34	39.8	1.5	1053	4	US-09-107-532A-3324	Sequence 3324, Ap
35	39.8	1.5	7721	3	US-08-772-270A-14	Sequence 14, Appl
36	39.8	1.5	1830121	4	US-09-557-884-1	Sequence 1, Appl
37	39.8	1.5	1830121	4	US-09-643-990A-1	Sequence 1, Appl
38	39	1.4	2590	4	US-09-598-401C-89	Sequence 89, Appl
39	39	1.4	3792	4	US-09-351-224E-10	Sequence 10, Appl
40	39	1.4	3792	4	US-09-677-488A-10	Sequence 10, Appl
41	39	1.4	3792	4	US-09-677-488B-10	Sequence 10, Appl
42	39	1.4	8093	4	US-10-204-708-32	Sequence 32, Appl
43	38.8	1.4	201	4	US-09-107-532A-3195	Sequence 3195, Ap
44	38.8	1.4	435	4	US-09-107-532A-3437	Sequence 3437, Ap
45	38.8	1.4	1149	4	US-09-134-001C-602	Sequence 602, App

## ALIGNMENTS

RESULT 1						
US-09-245-808-2						
Sequence 2, Application US/09245808						
Patent No. 6313277						
GENERAL INFORMATION:						
APPLICANT: Doyle, L. Austin						
APPLICANT: Abruzzo, Lynne V.						
TITLE OF INVENTION: Breast Cancer Resistance Protein (BCRP) and DNA which						
TITLE OF INVENTION: encodes it						
FILE REFERENCE: Ross UMD conversion						
CURRENT APPLICATION NUMBER: US/09/245, 808						
CURRENT FILING DATE: 1999-02-05						
EARLIER APPLICATION NUMBER: 60/073763						
EARLIER FILING DATE: 1998-02-05						
NUMBER OF SEQ ID NOS: 7						
SOFTWARE: Patentln Ver. 2.0						
SEQ ID NO 2						
LENGTH: 2418						
TYPE: DNA						
ORGANISM: Human MCF-7/AdrVP cells						
US-09-245-808-2						
Query Match						
Best Local Similarity 99.5%; Score 2352.8; DB 4; Length 2418;						
Matches 2360; Conservative 0; Mismatches 12; Indels 0; Gaps 0;						
QY	1	TTTAGAAGCGACCGTGCACATGCTTGGTGTCTTTGAAGTGAAGAACTGCTGTAGA	60			
DB	35	TTTAGAAGCGACCGTGCACATGCTTGGTGTCTTTGAAGTGAAGAACTGCTGTAGA	94			
QY	61	GTTTGTGGAAAGTCCGGGTGACTATCCCAACATTTAATCTTAAATGTTAAAGCGC	120			
DB	95	GTTTGTGGAAAGTCCGGGTGACTATCCCAACATTTAATCTTAAATGTTAAAGCGC	154			
QY	121	TGCTCCGAGCGACGATCTGAGCTTGTGAAGCCGAGCTCTATTAA	180			
DB	155	TGCTCCGAGCGACGATCTGAGCTTGTGAAGCCGAGCTCTATTAA	214			
QY	181	GCTGAAAGATTAATCTTCAGATGTTCCAGTAATGCGAAGTTTATCCCGTG	240			
DB	215	GCTGAAAGATTAATCTTCAGATGTTCCAGTAATGCGAAGTTTATCCCGTG	274			
QY	241	TCAAGAAACACCAATGCTTCCGCGACAGTTTCAATGACCTGAAGCATTTACT	300			
DB	275	TCAAGAAACACCAATGCTTCCGCGCGAGCTTCAATGACCTGAAGCATTTACT	334			
QY	301	GAAAGAGCTGTGAAGTTTCAATGATGATGATGATGATGATGATGATGATGATGAT	360			
DB	335	GAAAGAGCTGTGAAGTTTCAATGATGATGATGATGATGATGATGATGATGATGAT	394			

QY 361 CTACCTTGTGAAAAACAGTTGAGAAAGAAATATTATGAAATATCATGAGATCATGAAA 420  
DB 365 CTACCTTGTGAAAAACAGTTGAGAAAGAAATATTATGAAATATCATGAGATCATGAAA 454  
QY 421 CCGTGTCTCAACGCCATCTCTGGAGACCCACAGTGGAGGCAATCTTCCTTTATGATGTC 480  
DB 455 CCGTGTCTCAACGCCATCTCTGGAGACCCACAGTGGAGGCAATCTTCCTTTATGATGTC 514  
QY 481 TTAGCTGAGAGAAAGATCCAAAGTGAATTAATCTGAGATGTTCTGATTAATGAGACCG 540  
DB 515 TTAGCTGAGAGAAAGATCCAAAGTGAATTAATCTGAGATGTTCTGATTAATGAGACCG 574  
QY 541 CGACTGCGCAATTTCAAAATGTAATTCAGGTTACGTGTGTAACAAGATGATGTTGTGATGGCC 600  
DB 575 CGACTGCGCAATTTCAAAATGTAATTCAGGTTACGTGTGTAACAAGATGATGTTGTGATGGCC 634  
QY 601 ACTCTGAGCGTGGAGAGAAACCTTACAGTCTTCAGAGCTCTTGGGCTTGCAACAATATG 660  
DB 635 ACTCTGAGCGTGGAGAGAAACCTTACAGTCTTCAGAGCTCTTGGGCTTGCAACAATATG 694  
QY 661 ACGAATCTGAAAAAAGAAAGGATTAACAGGGTCAATGAAAGGTTAGTCTGAGATTA 720  
DB 695 ACGAATCTGAAAAAAGAAAGGATTAACAGGGTCAATGAAAGGTTAGTCTGAGATTA 754  
QY 721 GTGGAGAGCTCCAGAGTTGGAATCAGTTTATCCGTGTGTGTGTGAGAGAGAGAGAAA 780  
DB 755 GTGGAGAGCTCCAGAGTTGGAATCAGTTTATCCGTGTGTGTGTGAGAGAGAGAGAAA 814  
QY 781 AGGATTAATGAGAAATGAGAGGATTAATCACTGATCTTCATCTTGTCTGATGAGCT 840  
DB 815 AGGATTAATGAGAAATGAGAGGATTAATCACTGATCTTCATCTTGTCTGATGAGCT 874  
QY 841 ACAATGCTTGAATGAGATCAAGACAGCAAGCAATGCTGCTTGTCTGCTGAAAAGAGATGCT 900  
DB 875 ACAATGCTTGAATGAGATCAAGACAGCAAGCAATGCTGCTTGTCTGCTGAAAAGAGATGCT 934  
QY 901 AAGCAGGAGCAACATCATCTTCTCATATCATCAGCTCGATTAATTCATCTTCAAGTGG 960  
DB 935 AAGCAGGAGCAACATCATCTTCTCATATCATCAGCTCGATTAATTCATCTTCAAGTGG 994  
QY 961 TTTGATTAAGCTTCACTTTTGGCCCTCAGAGAACTTAATGTTCCACGGCCCTGCTCAGAG 1020  
DB 995 TTTGATTAAGCTTCACTTTTGGCCCTCAGAGAACTTAATGTTCCACGGCCCTGCTCAGAG 1054  
QY 1021 GCGTTGGAGATCTTTGAATCAAGCTGTGATTAATCACTGTGAGGCTTAATTAACCTGCAAG 1080  
DB 1055 GCGTTGGAGATCTTTGAATCAAGCTGTGATTAATCACTGTGAGGCTTAATTAACCTGCAAG 1114  
QY 1081 TTTCTTCTGAGATCATTAATGAGAAATTCACGTGTGCAATTAACAGAGAGAGAGAC 1140  
DB 1115 TTTCTTCTGAGATCATTAATGAGAAATTCACGTGTGCAATTAACAGAGAGAGAGAC 1174  
QY 1141 TTTAAAGCACAAGAGATCATAGAGCTTCAAGAGAGATTAAGCACTCATAGAAAAATTA 1200  
DB 1175 TTTAAAGCACAAGAGATCATAGAGCTTCAAGAGAGATTAAGCACTCATAGAAAAATTA 1234  
QY 1201 GGGAGATTTATGTAATCTCTCTCTTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1260  
DB 1235 GGGAGATTTATGTAATCTCTCTCTTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1294  
QY 1261 TCCGGGGGTGAGAGAGAGAGAGATCAAGCTTCAAGAGATCAAGCTCAACCACTCC 1320  
DB 1295 TCCGGGGGTGAGAGAGAGAGAGATCAAGCTTCAAGAGATCAAGCTCAACCACTCC 1354  
QY 1321 TTTCTGATCAACTCAAGATGAGGTTTCCAGAGGTTTATTCAAAAAATCTGCTGGATATCC 1380  
DB 1355 TTTCTGATCAACTCAAGATGAGGTTTCCAGAGGTTTATTCAAAAAATCTGCTGGATATCC 1414  
QY 1381 CAGGCTCTATAGCTCAAGATCTTGTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAG 1440  
DB 1415 CAGGCTCTATAGCTCAAGATCTTGTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAG 1474

QY 1441 TACTTGGGCTTAAAAATGATTTCTACTGGAATCCAGAAACAGAGCTGGGCTTCTCTTTC 1500  
DB 1475 TACTTGGGCTTAAAAATGATTTCTACTGGAATCCAGAAACAGAGCTGGGCTTCTCTTTC 1534  
QY 1501 CTGAGAGCAACACAGGTTTCCAGAGGTTTCCAGAGGTTTCCAGAGGTTTCCAGAGG 1560  
DB 1535 CTGAGAGCAACACAGGTTTCCAGAGGTTTCCAGAGGTTTCCAGAGGTTTCCAGAGG 1594  
QY 1561 AAGCTTTCATCATGATTAATCAATCAAGCGATTAATCAAGAGTCAATCAATTAATCCCTGGA 1620  
DB 1595 AAGCTTTCATCATGATTAATCAATCAAGCGATTAATCAAGAGTCAATCAATTAATCCCTGGA 1654  
QY 1621 AATCTGTTATGATTTATTAACCAATGAGAGATTAACCAAGATTAATTAATTAATCCCTGTA 1680  
DB 1655 AATCTGTTATGATTTATTAACCAATGAGAGATTAACCAAGATTAATTAATTAATCCCTGTA 1714  
QY 1681 GTGTACTTCATGTTAGATTAAGCAAGCAAGAGATGCTTCTCTGTTATGATGATTTAAC 1740  
DB 1715 GTGTACTTCATGTTAGATTAAGCAAGCAAGAGATGCTTCTCTGTTATGATGATTTAAC 1774  
QY 1741 CTATATAGTGTGCTTATTCAGCAGTTCATGAGCACTGAGCATATGACAGAGTCAAGAT 1800  
DB 1775 CTATATAGTGTGCTTATTCAGCAGTTCATGAGCACTGAGCATATGAGCATATGAGAT 1834  
QY 1801 GTGTGTTCTGAGCAACATCTCATGACCATCTGTTGTGTTATGATGATTTTTC 1860  
DB 1835 GTGTGTTCTGAGCAACATCTCATGACCATCTGTTGTGTTATGATGATTTTTC 1894  
QY 1861 GGTCTGTGTGATATCTCAACCAATGATGATCTTGGCTGTCAATGCTTCAATGATTCAGC 1920  
DB 1895 GGTCTGTGTGATATCTCAACCAATGATGATCTTGGCTGTCAATGCTTCAATGATTCAGC 1954  
QY 1921 ATTCCAGATATGAGATTTACAGGCTTGGAGCAATTAATGATTTTGGAGCAAACTTCTGC 1980  
DB 1955 ATTCCAGATATGAGATTTACAGGCTTGGAGCAATTAATGATTTTGGAGCAAACTTCTGC 2014  
QY 1981 CAGAGCTCAATGCAACAGAGAAACAACTCTTGAATGATGCAACATGTAATGAGAGAA 2040  
DB 2015 CAGAGCTCAATGCAACAGAGAAACAACTCTTGAATGATGCAACATGTAATGAGAGAA 2074  
QY 2041 TATTTGTTAAGAGGAGCTGATCTGCAACCTGGGCTTGTGAAAGATCAAGCTGGCC 2100  
DB 2075 TATTTGTTAAGAGGAGCTGATCTGCAACCTGGGCTTGTGAAAGATCAAGCTGGCC 2134  
QY 2101 TTGGCTTGTATGATTTTATTTTCTCACAATGCTGCAATGCTGAAATGTTATTTCTTAA 2160  
DB 2135 TTGGCTTGTATGATTTTATTTTCTCACAATGCTGCAATGCTGAAATGTTATTTCTTAA 2194  
QY 2161 AATATTTCTTAATTTCCCTTAATTCAGTATGATTTATCCCTCAATTAAGAGAGAGAG 2220  
DB 2195 AATATTTCTTAATTTCCCTTAATTCAGTATGATTTATCCCTCAATTAAGAGAGAGAG 2254  
QY 2221 TTGATTTGAAGTATCAATCAAGTTTGTGTTTGTGTTTCTGCTTCCCTGCAATCACTG 2280  
DB 2255 TTGATTTGAAGTATCAATCAAGTTTGTGTTTGTGTTTCTGCTTCCCTGCAATCACTG 2314  
QY 2281 TTGACAGAGAGCAATGTTTAAAGAGATCAATTTTAAAGATCAACAACTGAAATTA 2340  
DB 2315 TTGACAGAGAGCAATGTTTAAAGAGATCAATTTTAAAGATCAACAACTGAAATTA 2374  
QY 2341 AACATGAAAGAACCCAGACATCATGATCGC 2372  
DB 2375 AACATGAAAGAACCCAGACATCATGATCAC 2406

RESULT 2  
US-09-245-808-7  
; Sequence 7, Application US/09245808  
; Patent No. 6313277  
; GENERAL INFORMATION:  
; APPLICANT: Doyle, L. Austin  
; APPLICANT: Abruzzo, Lynne V.  
; APPLICANT: Ross, Douglas D.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 18, 2004, 13:01:06 ; Search time 4551 Seconds

(without alignments)  
17841.193 Million cell updates/sec

Title: US-09-856-927-1

Perfect score: 2719

Sequence: 1 tttaggaacgcacgtgcac.....catlaagtaataagaactt 2719

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
EST:  
1: em\_estda:\*  
2: em\_esthm:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estcom:\*  
17: em\_gse\_hum:\*  
18: em\_gse\_inv:\*  
19: em\_gse\_pin:\*  
20: em\_gse\_vrt:\*  
21: em\_gse\_fun:\*  
22: em\_gse\_mam:\*  
23: em\_gse\_mus:\*  
24: em\_gse\_pro:\*  
25: em\_gse\_rtd:\*  
26: em\_gse\_phg:\*  
27: em\_gse\_vr1:\*  
28: gb\_gsal1:\*  
29: gb\_gsal2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	982.6	36.1	1345	11	U66681 Homo sapien
2	970.2	35.7	2869	11	AK030831 Mus muscu
3	970.2	35.7	3001	11	AK051880 Mus muscu
4	913	33.6	1034	12	BMS44411 AGENCOURT

5	885.4	32.6	925	13	EX370114	EX370114	EX370114
6	859.8	31.6	1201	13 <td>EX382704</td> <td>EX382704</td> <td>EX382704</td>	EX382704	EX382704	EX382704
7	847.2	31.2	907	13 <td>EX329115</td> <td>EX329115</td> <td>EX329115</td>	EX329115	EX329115	EX329115
8	821.2	30.2	888	13 <td>EX454801</td> <td>EX454801</td> <td>EX454801</td>	EX454801	EX454801	EX454801
9	816.4	30.0	883	9	AU118354	AU118354	AU118354
10	782.6	28.8	1201	13 <td>EX382703</td> <td>EX382703</td> <td>EX382703</td>	EX382703	EX382703	EX382703
11	773	28.4	839	12 <td>BG573081</td> <td>BG573081</td> <td>BG573081</td>	BG573081	BG573081	BG573081
12	753.8	27.7	821	12 <td>BI918550</td> <td>BI918550</td> <td>BI918550</td>	BI918550	BI918550	BI918550
13	725.6	26.7	794	10 <td>AM155343</td> <td>AM155343</td> <td>AM155343</td>	AM155343	AM155343	AM155343
14	714	26.3	740	12 <td>BG573904</td> <td>BG573904</td> <td>BG573904</td>	BG573904	BG573904	BG573904
15	679.8	25.0	727	9	AU137863	AU137863	AU137863
16	679.4	25.0	836	14 <td>CF593368</td> <td>CF593368</td> <td>CF593368</td>	CF593368	CF593368	CF593368
17	678.6	25.0	961	12 <td>BG386140</td> <td>BG386140</td> <td>BG386140</td>	BG386140	BG386140	BG386140
18	668.6	24.6	739	11 <td>BC011032</td> <td>BC011032</td> <td>BC011032</td>	BC011032	BC011032	BC011032
19	666.2	24.5	708	14 <td>CD355766</td> <td>CD355766</td> <td>CD355766</td>	CD355766	CD355766	CD355766
20	641.2	23.6	666	9	AL135203	AL135203	AL135203
21	637.4	23.4	697	9	AV728650	AV728650	AV728650
22	632	23.2	879	12 <td>BM423310</td> <td>BM423310</td> <td>BM423310</td>	BM423310	BM423310	BM423310
23	624.8	23.0	631	12 <td>BI497116</td> <td>BI497116</td> <td>BI497116</td>	BI497116	BI497116	BI497116
24	614.4	22.6	950	13 <td>EX370198</td> <td>EX370198</td> <td>EX370198</td>	EX370198	EX370198	EX370198
25	611.2	22.5	620	9	AL603604	AL603604	AL603604
26	608.4	22.4	787	13 <td>EX418619</td> <td>EX418619</td> <td>EX418619</td>	EX418619	EX418619	EX418619
27	607	22.3	671	12 <td>BI086640</td> <td>BI086640</td> <td>BI086640</td>	BI086640	BI086640	BI086640
28	592.6	21.8	664	14 <td>BG436056</td> <td>BG436056</td> <td>BG436056</td>	BG436056	BG436056	BG436056
29	586.6	21.6	644	14 <td>CB552639</td> <td>CB552639</td> <td>CB552639</td>	CB552639	CB552639	CB552639
30	584	21.5	954	13 <td>EX370199</td> <td>EX370199</td> <td>EX370199</td>	EX370199	EX370199	EX370199
31	567.4	20.9	588	10 <td>BE048216</td> <td>BE048216</td> <td>BE048216</td>	BE048216	BE048216	BE048216
32	562.4	20.7	607	12 <td>BI553320</td> <td>BI553320</td> <td>BI553320</td>	BI553320	BI553320	BI553320
33	558.2	20.5	712	12 <td>BG689713</td> <td>BG689713</td> <td>BG689713</td>	BG689713	BG689713	BG689713
34	558	20.5	873	10 <td>BE207763</td> <td>BE207763</td> <td>BE207763</td>	BE207763	BE207763	BE207763
35	529.8	19.5	907	13 <td>EX409712</td> <td>EX409712</td> <td>EX409712</td>	EX409712	EX409712	EX409712
36	518.6	19.1	653	10 <td>BF674699</td> <td>BF674699</td> <td>BF674699</td>	BF674699	BF674699	BF674699
37	506.8	18.6	672	14 <td>CB424319</td> <td>CB424319</td> <td>CB424319</td>	CB424319	CB424319	CB424319
38	501.4	18.4	594	9	AU157347	AU157347	AU157347
39	500.4	18.4	534	9	AU144940	AU144940	AU144940
40	495	18.2	586	12 <td>BG705688</td> <td>BG705688</td> <td>BG705688</td>	BG705688	BG705688	BG705688
41	493.2	18.1	498	9	AL872416	AL872416	AL872416
42	492.4	18.1	722	14 <td>CB959914</td> <td>CB959914</td> <td>CB959914</td>	CB959914	CB959914	CB959914
43	492	18.1	492	10 <td>AM270740</td> <td>AM270740</td> <td>AM270740</td>	AM270740	AM270740	AM270740
44	478	17.6	489	9	AA933960	AA933960	AA933960
45	473.4	17.4	475	10 <td>BE672952</td> <td>BE672952</td> <td>BE672952</td>	BE672952	BE672952	BE672952

## ALIGNMENTS

RESULT 1  
LOCUS U66681 1345 bp mRNA linear HTC 23-JUL-2001  
DEFINITION Homo sapiens clone EST157481 mRNA sequence.  
ACCESSION U66681  
VERSION U66681.1 GI:1906566  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 1345)  
Allikmetz, R., Gerard, B., Hutchinson, A., and Dean, M.  
Characterization of the human ABC superfamily: isolation and mapping of 21 new genes using the expressed sequence tags database Hum. Mol. Genet. 5 (10), 1649-1655 (1996)  
JOURNAL MEDLINE  
MEDLINE  
PUBMED 8894702  
REFERENCE  
2 (bases 1 to 1345)  
Allikmetz, R., Gerard, B., and Dean, M.  
Direct Submission  
Submitted (12-AUG-1996) Human Genetics Section, National Cancer Institute, NCI-PCRD, Frederick, MD 21702, USA  
FEATURES  
source  
1..1345  
/organism="Homo sapiens"  
/mol\_type="mRNA"

```

/d_b_xref="taxon:9606"
/clone="EST157481"
/note="similar to ATP-binding cassette transporter"

```

Query Match	36.1%	Score	982.6	DB	11	Length	1345
Best Local Similarity	82.1%	Pred.	No. 2.4e-226				
Matches	1288	Conservative	0	Mismatches	55	Indels	226
						Gaps	5

QY	1153	GAGATCATAGAGCCCTTCCAGCAGAGATAGACCATCATAGAAAAATTGCGGAGATTAT	1212
Db	1	GAGATCATAGAGCCCTTCCAGCAGAGATAGACCATCATAGAAAAATTGCGGAGATTAT	60
QY	1213	GTCAACTCTCTCTTCTACAAAGAGCAAAAGTGAATTACATCACTTCCGGGGGTGAG	1272
Db	61	GTCAACTCTCTCTTCTACAAAGAGCAAAAGTGAATTACATCACTTCCGGGGGTGAG	120
QY	1273	AAGAAGAAAGATACACAGTCTTCAAGGAGATCAGCTACACCCTCTGTGATGAA	1332
Db	121	AAGAAGAAAGATACACAGTCTTCAAGGAGATCAGCTACACCCTCTGTGATGAA	180
QY	1333	CTCAGATGGGTTTCCAGCGTTCAATTCAAAACTTGTGGTAATCCCAAGGCTCTATA	1392
Db	181	CTCAGATGGGTTTCCAGCGTTCAATTCAAAACTTGTGGTAATCCCAAGGCTCTATA	240
QY	1393	GCTCAGATCATGTGACAGTGGTACTGGGACTGGGTATATAGTGGCAATTACTTGGGGCTA	1452
Db	241	GCTCAGATCATGTGACAGTGGTACTGGGACTGGGTATATAGTGGCAATTACTTGGGGCTA	300
QY	1453	AAAAATGATTTCTACTGGAATCCAGAACAGAGCTGGGGTTCTTCTTCTCTGACGACAAAC	1512
Db	301	AAAAATGATTTCTACTGGAATCCAGAACAGAGCTGGGGTTCTTCTTCTCTGACGACAAAC	360
QY	1513	CAGTGTTCAGAGAGTTCACGCCGTGGAACCTTTGTGTGTAGAAAGAAAGCTCTTCATA	1572
Db	361	CAGTGTTCAGAGAGTTCACGCCGTGGAACCTTTGTGTGTAGAAAGAAAGCTCTTCATA	420
QY	1573	CATGAATACATCAGCGGATACTACAGAGTGTATCTTATTTCTCTGAAAACTGTATCT	1632
Db	421	CATGAATACATCAGCGGATACTACAGAGTGTATCTTATTTCTCTGAAAACTGTATCT	480
QY	1633	GATTATATACCATAGAGATGTTACCAAGTATTATTTACTGTATAGTGTACTTCATG	1692
Db	481	GATTATATACCATAGAGATGTTACCAAGTATTATTTACTGTATAGTGTACTTCATG	540
QY	1693	TTAGAGTTGAAGCCAAAGGACAGATGCGCTTCCTCGTATATGTTTAACTTATAGTGGTG	1752
Db	541	TTAGAGTTGAAGCCAAAGGACAGATGCGCTTCCTCGTATATGTTTAACTTATAGTGGTG	600
QY	1753	GCTTATTCAGCCAGTTCATGGAOCTGGCCATAGACAGCAGTCAAGTGTGGTTCTGTA	1812
Db	601	GCTTATTCAGCCAGTTCATGGAOCTGGCCATAGACAGCAGTGTGGTTCTCTGTA	660
QY	1813	GGAACACTCTCATAGACCATCTGTTTGTGTTATGATGATTTTTCAAGTGTGTGGTC	1872
Db	661	GGAACACTCTCATAGACCATCTGTTTGTGTTATGATGATTTTTCAAGTGTGTGGTC	720
QY	1873	AATCTCAACAACATTGCATCTT--GGCTGTCAAT--GGCTTCAGTACTTCAGCATTCACGAT	1932
Db	721	AATCTCAACAACATTGCATCTTGGGCTGTCAATGGGCTTCAGTACTTCAGCATTCACG--	778
QY	1931	ATGAGTTTACGGCTTTGACAGCATTAATGAATTTTGGGACAAAACTTCGGCCAGAGATCA	1991
Db	779	-----GATATGGGATTTTACGGGCTTTGACAGCATTA	809
QY	1991	ATGCAACAGAAAAACATCCTTGTAACTATGCAACATGTACTGGCAGAAAGATATTGTGTAA	2051
Db	810	ATGAA-----TTT	817
QY	2051	AGCAGAGCATGCATCTCTCAACCTGGGGGCTGTGTGAAAGAAATCAGTGGGCTTGCTGTGA	2111
Db	818	TTGGGGGACAAAACTTCTGCCCCCGGGGCTTCAATGCGAACCAAGGGGAA-----	866

QY	2111	TGATGTGTAATTCCTGCACATGCCCTACCTGAAATGTATATCTTAAAAAATATCTT	217
Db	867	-----	866
QY	2171	AAATTCCTTAATTCAGATGATTTATCTCTACATAAAAAGAGACCTTGATTGAA	2230
Db	867	-----	866
QY	2231	GTATTCATCAAGTTTTTTTGTTGTTCTTCGTTCCCTGCATCACACTGTGCACACA	2290
Db	867	-----ACNNNTTTTGTGTGTTTCTGTTCCCTGGCATCACACTGTGCACACA	916
QY	2281	GCAATGTTTTTAAAGATATACATTTTAAAGATACAGACAAACGATTTAAACATGAAG	2350
Db	917	GCAATGTTTTTAAAGATATACATTTTAAAGATACAGACAAACGATTTAAACATGAAG	976
QY	2351	AACCCAGACATCATGTATCGCATATTAGTTAATCTCTCTCAGACAGTAAACCATGGGAG	2410
Db	977	AACCCAGACATCATGTATCGCATATTAGTTAATCTCTCTCAGACAGTAAACCATGGGAG	1036
QY	2411	AAATCGGTCAATTTTATCTTAACTAAAAAGAGAAATGAAATTCGTGAAATCCTGCACA	2470
Db	1037	AAATCGGTCAATTTATTAATCTTAAAAAGAGAAATGAAATTCGTGAAATCCTGCACA	1096
QY	2471	GTTATTAATGTCCTGCGCATTTGTTTCTCTCATCTTTAAATGAATAGGTAGTTAGTAC	2530
Db	1097	GTTATTAATGTCCTGCGCATTTGTTTCTCTCATCTTTAAATGAATAGGTAGTTAGTAC	1156
QY	2531	CCTTCAGCTTAATCACTTATGATGCTATGCTTGGCATTTTAAATATATGACAAATGT	2590
Db	1157	CCTTCAGCTTAATCACTTATGATGCTATGCTTGGCATTTTAAATATATGACAAATGT	1216
QY	2591	ATTATGCTATTAATCGAAATGTAAATTTGAAATATGTTGAAAAAAGATTCTGTCTTAT	2656
Db	1217	ATTATGCTATTAATCGAAATGTAAATTTGAAATATGTTGAAAAAAGATTCTGTCTTAT	1276
QY	2651	AGGGTAAAAAAGCAGCCGCTGATAGAAAAAATCTTTTGATAGACACATTTAAAGTTA	2710
Db	1277	AGGGTAAAAAAGCAGCCGCTGATAGAAAAAATCTTTTGATAGACACATTTAAAGTTA	1336
QY	2711	ATAAGACTT 2719	
Db	1337	ATAAGACTT 1345	

AK030831	LOCUS	AK030831	2869 bp	mRNA	linear	HTC 18-SEP-2003
DEFINITION		Mus musculus adult male thymus cDNA, RIKEN full-length enriched library, clone:5830416h02 product:ATP-binding cassette, sub-family G (WHITE), member 3, full insert sequence.				
ACCESSION		AK030831				
VERSION		AK030831.1	GI:26082008			
KEYWORDS		HTC, CAP trapper				
SOURCE		Mus musculus (house mouse)				
ORGANISM		Mus musculus				
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Eukaryota; Eutheria; Rodentia; Sciurgnathii; Muridae; Murinae; Mus.				
REFERENCE		1	Carninci, P. and Hayashizaki, Y.			
AUTHORS			High-efficiency full-length cDNA cloning			
TITLE			Meth. Enzymol. 303, 19-44 (1999)			
JOURNAL			99279253			
MEDLINE			10349636			
PUBMED						
REFERENCE		2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
AUTHORS			Normalization and subtraction of cap-trapper-selected cDNAs to			
TITLE			prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL			Genome Res. 10 (10), 1617-1630 (2000)			
MEDLINE			20499374			
PUBMED			11042159			
REFERENCE		3				



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 18, 2004, 11:16:26 ; Search time 6989 Seconds

(without alignments)  
16862.170 Million cell updates/sec

Title: US-09-856-927-1  
Perfect score: 2719  
Sequence: 1 tttaggaacgcacgcgtgcac.....cattaagtcatacgaactt 2719

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:\*

1: gb\_ba:\*

2: gb\_hcg:\*

3: gb\_in:\*

4: gb\_cm:\*

5: gb\_cv:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pt:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vl:\*

15: em\_pa:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_un:\*

29: em\_vl:\*

30: em\_hcg\_hum:\*

31: em\_hcg\_inv:\*

32: em\_hcg\_other:\*

33: em\_hcg\_mus:\*

34: em\_hcg\_pln:\*

35: em\_hcg\_rod:\*

36: em\_hcg\_mam:\*

37: em\_hcg\_vrt:\*

38: em\_ey:\*

39: em\_hcgo\_hum:\*

40: em\_hcgo\_mus:\*

41: em\_hcgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2719	100.0	2719	6	BD234217
2	2719	100.0	2719	6	AX322795
3	2719	100.0	2719	6	AF103796
4	2354.6	86.6	2574	6	AX061220
5	2354.4	86.6	2418	6	AF098951
6	2352.8	86.5	2418	6	AF177881
7	2352.8	86.5	2418	6	BD124164
8	2332.4	82.1	2547	6	BC021281
9	2221.4	81.7	2247	6	AX322812
10	2221.4	81.7	2247	9	AY017168
11	2212	81.4	2212	9	AY289766
12	2090.4	76.9	2788	6	BD192221
13	2078.6	76.4	2085	9	AF463519
14	2044.2	75.2	2077	6	AX878056
15	2044.2	75.2	2077	6	BD157000
16	2044.2	75.2	2077	9	AK002040
17	2022.2	74.4	2027	6	BD167779
18	2022.2	74.4	2027	6	AB051855
19	2010	73.9	2053	6	BD167789
20	1963.2	72.2	1968	6	BD133828
21	1963.2	72.2	1968	9	AB056867
22	1960	72.1	1998	6	BD133830
23	1955.2	71.9	1998	6	BD133829
24	1866.8	62.0	1690	9	AY288307
25	1506.4	53.4	2112	4	SSC420927
26	1431.8	52.7	1469	9	AF093771
27	1408	51.8	2175	10	AY089397
28	1403.2	51.6	2139	10	AY089396
29	1400	51.5	2130	10	AY089398
30	1398.8	51.4	2083	10	AB105817
31	1393	51.2	2520	10	BC053730
32	1392.2	51.2	2025	6	AX322799
33	1392.2	51.2	2025	10	AF140218
34	1363	50.1	1974	10	AB094089
35	1303.6	47.9	176392	2	AC135954
36	1240	45.6	187712	9	AC027088
37	1103.6	40.6	1144	9	AF093772
38	971.8	35.7	2237	10	AF324242
39	746.4	27.5	795	6	AR177886
40	746.4	27.5	795	6	BD124169
41	682	25.1	181576	9	AC084732
42	679.8	25.0	727	6	AX869789
43	679.8	25.0	727	6	BD149851
44	574.6	21.1	231679	2	AC123319
45	574.6	21.1	255881	2	AC099252

## ALIGNMENTS

RESULT 1  
BD234217  
LOCUS  
DEFINITION  
BD234217  
ATP-binding cassette genes and proteins for diagnosis and remedy of  
lipid disorders and inflammatory diseases.  
ACCESSION  
BD234217  
VERSION  
BD234217.1 GI:33043587  
KEYWORDS  
JP 2002525111-A/27.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 2719)  
Schmitz/G. and Klucken,J.  
TITLE  
ATP-binding cassette genes and proteins for diagnosis and remedy of

JOURNAL

Lipid disorders and inflammatory diseases  
Patent: JP 2002525111-A 27 13-AUG-2002;

COMMENT

BAYER AKTIENGESELLSCHAFT  
OS Homo sapiens (human)  
PN JP 2002525111-A/27  
PD 13-AUG-2002 JP 2000572359  
PR 21-SEP-1998 US 60/101706  
PI GEND SCHMITZ, JOCHEN KLUCKEN  
PC C12N15/09, A61K38/00, A61K45/00, A61K48/00, A61P3/06, A61P9/10, PC  
A61P29/00,  
PC C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC  
C12P21/02, C1201/68,  
PC G01N33/53 C12N15/00, C12N5/00, A61K37/02  
CC human cDNA of ABCG2  
FT Key Location/Qualifiers  
FT source 1. 2719  
FT Location/Qualifiers  
1. 2719 /organism='Homo sapiens (human)'.  
/mol\_type='genomic DNA'  
/db\_xref='taxon:9606'

FEATURES  
source

ORIGIN

Query Match 100.0%; Score 2719; DB 6; Length 2719;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2719; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGAGAGGCGACCGTGCACATGCTTGGTGTCTTTGTTAGTGAAGAACTGCTGCTTTAGA 60  
DB 1 TTGAGAGGCGACCGTGCACATGCTTGGTGTCTTTGTTAGTGAAGAACTGCTGCTTTAGA 60  
QY 61 GTTGTGTTGAGAGGTCCGGGTGACTCATCCCAACATTTACATCTTAATTGTTAAAGCGC 120  
DB 61 GTTGTGTTGAGAGGTCCGGGTGACTCATCCCAACATTTACATCTTAATTGTTAAAGCGC 120  
QY 121 TGGCTCCGAGCGGACGATCTGTAGATCTGTAGCCTTTGGTTAAGACCGAGCTCTATTAA 180  
DB 121 TGGCTCCGAGCGGACGATCTGTAGATCTGTAGCCTTTGGTTAAGACCGAGCTCTATTAA 180  
QY 121 TGGCTCCGAGCGGACGATCTGTAGATCTGTAGCCTTTGGTTAAGACCGAGCTCTATTAA 180  
DB 121 TGGCTCCGAGCGGACGATCTGTAGATCTGTAGCCTTTGGTTAAGACCGAGCTCTATTAA 180  
QY 181 GGTGAAAGATGAAAGTCTCCAGATGTCTCCAGTAATGTGAAAGTTTATCCCAAGT 240  
DB 181 GGTGAAAGATGAAAGTCTCCAGATGTCTCCAGTAATGTGAAAGTTTATCCCAAGT 240  
QY 181 GGTGAAAGATGAAAGTCTCCAGATGTCTCCAGTAATGTGAAAGTTTATCCCAAGT 240  
DB 181 GGTGAAAGATGAAAGTCTCCAGATGTCTCCAGTAATGTGAAAGTTTATCCCAAGT 240  
QY 241 TCAAGAGAAACCAATGAGTCTCCGCGACAGTTTCAATGAGCTGAAAGCAATTTACT 300  
DB 241 TCAAGAGAAACCAATGAGTCTCCGCGACAGTTTCAATGAGCTGAAAGCAATTTACT 300  
QY 301 GAAGAGCTGTGTTAGTTTCAATCACTGTCTATCGAGTAAACTGAAAGAGTGGCTT 360  
DB 301 GAAGAGCTGTGTTAGTTTCAATCACTGTCTATCGAGTAAACTGAAAGAGTGGCTT 360  
QY 361 CTACCTTGTGAAAGCAAGTGAAGAAATATTCGAATGATCATGAGATCATGAGAA 420  
DB 361 CTACCTTGTGAAAGCAAGTGAAGAAATATTCGAATGATCATGAGATCATGAGAA 420  
QY 421 CCGGTCTGAAGCCATCTGCGGACCCAGAGTGAAGGCAATCTTCGTTATTAATGTC 480  
DB 421 CCGGTCTGAAGCCATCTGCGGACCCAGAGTGAAGGCAATCTTCGTTATTAATGTC 480  
QY 481 TTAGCTCAAGAGAAATCATGATGATGATGATGATGATGATGATGATGATGATGATG 540  
DB 481 TTAGCTCAAGAGAAATCATGATGATGATGATGATGATGATGATGATGATGATGATG 540  
QY 541 CGACCTGCAATTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
DB 541 CGACCTGCAATTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
QY 601 ACTGTACGAGTGAAGAAAGTCAAGTCTGAGAGTCTTCCGCTTGAACAATGATG 660  
DB 601 ACTGTACGAGTGAAGAAAGTCAAGTCTGAGAGTCTTCCGCTTGAACAATGATG 660

QY 661 ACGATCATGAAAAAGAAAGCGATTAACAGGCTCATTTGAAGATGATGATGATGATGATGAT 720  
DB 661 ACGATCATGAAAAAGAAAGCGATTAACAGGCTCATTTGAAGATGATGATGATGATGATGAT 720  
QY 721 GTGGCAGACTCCAGGTTGGAACTCACTTATTCCTGTGTGTGTGTGTGTGTGTGTGTGTGT 780  
DB 721 GTGGCAGACTCCAGGTTGGAACTCACTTATTCCTGTGTGTGTGTGTGTGTGTGTGTGTGT 780  
QY 781 AGGACTAGTATGAGAAATGAGGCTTATCATGATGATGATGATGATGATGATGATGATGATG 840  
DB 781 AGGACTAGTATGAGAAATGAGGCTTATCATGATGATGATGATGATGATGATGATGATGATG 840  
QY 841 ACAATGCTTGAAGCTCAAGACAGCAAAAGCTGTCTTTTGTCTCTGAAAGAGATGCT 900  
DB 841 ACAATGCTTGAAGCTCAAGACAGCAAAAGCTGTCTTTTGTCTCTGAAAGAGATGCT 900  
QY 901 AAGCAGGAGCAAAATCATCTTCTCATTCATGAGCTGATGATGATGATGATGATGATGATG 960  
DB 901 AAGCAGGAGCAAAATCATCTTCTCATTCATGAGCTGATGATGATGATGATGATGATGATG 960  
QY 961 TTGATAGCTTCACTTATTTGGCTTCAAGAACTTATGTTTCAAGGAGCTGCTCAGAG 1020  
DB 961 TTGATAGCTTCACTTATTTGGCTTCAAGAACTTATGTTTCAAGGAGCTGCTCAGAG 1020  
QY 1021 GCTTGGGATCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080  
DB 1021 GCTTGGGATCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080  
QY 1081 TTCTTCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140  
DB 1081 TTCTTCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140  
QY 1141 TTAAAGCAGAGATCATAGAGCTTCCAGAGAGTAAAGCATGATGATGATGATGATGATGATG 1200  
DB 1141 TTAAAGCAGAGATCATAGAGCTTCCAGAGAGTAAAGCATGATGATGATGATGATGATGATG 1200  
QY 1261 TCCGGGGGAGAGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320  
DB 1261 TCCGGGGGAGAGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320  
QY 1321 TTCTGTATCACTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380  
DB 1321 TTCTGTATCACTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380  
QY 1381 CAGGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440  
DB 1381 CAGGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440  
QY 1441 TACTTGGGCTTAAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500  
DB 1441 TACTTGGGCTTAAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500  
QY 1501 CTGACGACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560  
DB 1501 CTGACGACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560  
QY 1561 AAGCTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620  
DB 1561 AAGCTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620  
QY 1621 AAATCTTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680  
DB 1621 AAATCTTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680  
QY 1681 GTGTACTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740  
DB 1681 GTGTACTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740  
QY 1741 CTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800